

USING ARTIFICIAL INTELLIGENCE TO DIGITIZE AND ANALYZE PARLIAMENTARY BILLS IN SUB-SAHARAN AFRICA - 1ST AFRICANLP WORKSHOP, INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS (ICLR 2020)

Adewale Akinfaderin;

;

© 2021, ADEWALE AKINFADERIN



This work is licensed under the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted use, distribution, and reproduction, provided the original work is properly credited.

Cette œuvre est mise à disposition selon les termes de la licence Creative Commons Attribution (<https://creativecommons.org/licenses/by/4.0/legalcode>), qui permet l'utilisation, la distribution et la reproduction sans restriction, pourvu que le mérite de la création originale soit adéquatement reconnu.

Using Artificial Intelligence to Digitize and Analyze Parliamentary Bills in Sub-Saharan Africa

Adewale Akinfaderin

1st AfricaNLP Workshop,
International Conference on Learning Representations (ICLR 2020)

We Built NASS-AI



Dataset - Samples of Parliamentary Bills from Nigeria

SB 348

A BILL

FOR

AN ACT TO REPEAL THE NATIONAL ECONOMIC RECONSTRUCTION FUND ACT AND TO TRANSFER THE ASSETS AND LIABILITIES TO THE BANK OF INDUSTRY

Commencement

BE IT ENACTED by the National Assembly of the Federal Republic of Nigeria as follows:

1. The National Economic Reconstruction Fund Act is hereby repealed and accordingly, the National Economic Reconstruction Fund established by it is hereby dissolved.

2. (1) The assets, funds resources and other movable or immovable property which before the commencement of this Act were vested in the National Economic Reconstruction Fund (in this Act referred to as "the Fund") are hereby vested in the Bank of Industry (in this Act referred to as "the Bank").

(2) The rights, interests, obligations and liabilities of the Fund under any contract or instrument whether at law or in equity apart from any contract or instrument, shall by virtue of this Act be assigned to and vested in the Bank.

(3) Any contract, or instrument as mentioned in subsection (2) of this section, shall be of the same force and effect against or in favour of the Bank and shall be enforceable as fully and effectively as if, instead of the Fund, the Bank had been named in the contract or instrument or had been a party to the contract or instrument.

(4) The Bank shall be subject to all obligations and liabilities to which the Fund was subject before the commencement of this Act and all other persons shall have the same rights, powers and remedies against the Bank as they had against the Fund before the commencement of this Act.

3. Any proceedings or cause of action pending or existing before the commencement of this Act on which could have been enforced by or taken against the Fund in respect of any right, interest, obligation or liability of the Fund may be amended, continued enforced by or taken against the Bank as if the Bank and not the Fund had been a party thereto.

Repeal Cap 254
LFN

Transfer of assets, etc. to
the Bank of industry
Limited

Proceedings, etc.
of the Fund.

A bill to regulate local government elections in Nigeria

[SB, 215]

A BILL

FOR

AN ACT TO PROHIBIT THE USE OF LIFE BULLETS OR THE NIGERIAN ARMY TO QUEL STUDENTS PROTESTS AND ALL CIVIL PROTESTS OF THE NIGERIAN LABOUR CONGRESS

Sponsored by SENATOR TIMOTHY AKE ADIGUN
(Member of the Senate)

BE IT ENACTED by the National Assembly of the Federal Republic of Nigeria as follows:—

1. Notwithstanding anything to the contrary contained in any enactment, law or Provisions however contained in this Act shall apply to restrict the deployment of soldiers to quell students' riots and all civil protests of the Nigerian Labour Congress.

2. The Government of the Federation or of a state shall not deploy soldiers to quell students protests and any demonstrations or protests of the Nigerian Labour Congress.

3. The use of life bullets to avert, stop or quell students and Nigerian Labour Congress protests is hereby prohibited.

4. The responsibility of averting or preventing or quelling any protests organized and embarked upon by students and the Nigerian Labour Congress shall be solely conducted by the Police in conjunction and conformity with riots the Nigeria Police Act CNP 359 Laws of the Federation of Nigeria 1990 or such other law in existence.

5.—(1) In the likely event that life bullets are to be used to quell such protest, it must by a resolution of the National Assembly on the request of the Inspector-General of Police or any other person acting on his behalf by way of application stating out in details the grounds as basis for the grant.

(2) An applicant on under Section 5(i) hereof shall be considered by the Senate within twenty four hours on receipt when the Senate is not on recess.

Commence-
ment.

Application of
the Bill.

Restriction of
life bullets.

Prohibition of
the use of life
bullets.

Responsibility
of quelling.

Exceptions
and time limit

A bill to prohibit the use of life bullets and military to quell protests

[SB, 330]

A BILL

FOR

OTHER MATTERS CONNECTED THEREWITH

Sponsored by SENATOR ROBERT B. KOLAWOLE
(Member of the Senate)

BE IT ENACTED by the National Assembly of the Federal Republic of Nigeria as follows:

1.—(1) As from the commencement of this Act, and subject to the provisions thereto, every other person in Nigeria shall be screened and be treated of cancer and shall—
(a) be required to register with any teaching hospital or medical centre (hereafter in this Act referred to as "Hospital") nearest to her place of abode, within two weeks of diagnosis;

(b) upon compliance with the requirement of (a) of this Subsection, be qualified to receive free medical treatment and medical services.
2.—(1) The Medical Director of the hospital or any other person Acting in that behalf to which this Act relates shall—

(a) Not later than one week following the screening and registration in his/her hospital of cancer patient(s), transmit all the relevant information on particularly with respect to any such patient(s) to the relevant government authority or Ministry charged with the responsibility for health management Services;

(b) Not refuse the screening, registration and treatment of such patients; and
(c) Ensure the provision of free treatment and other health services as specified in section 1 (4).

3.—(1) The Government of the Federation shall—
(a) Provide screening and treatment Equipment for all Hospital

Directives as
to free
treatment and
medical
services.

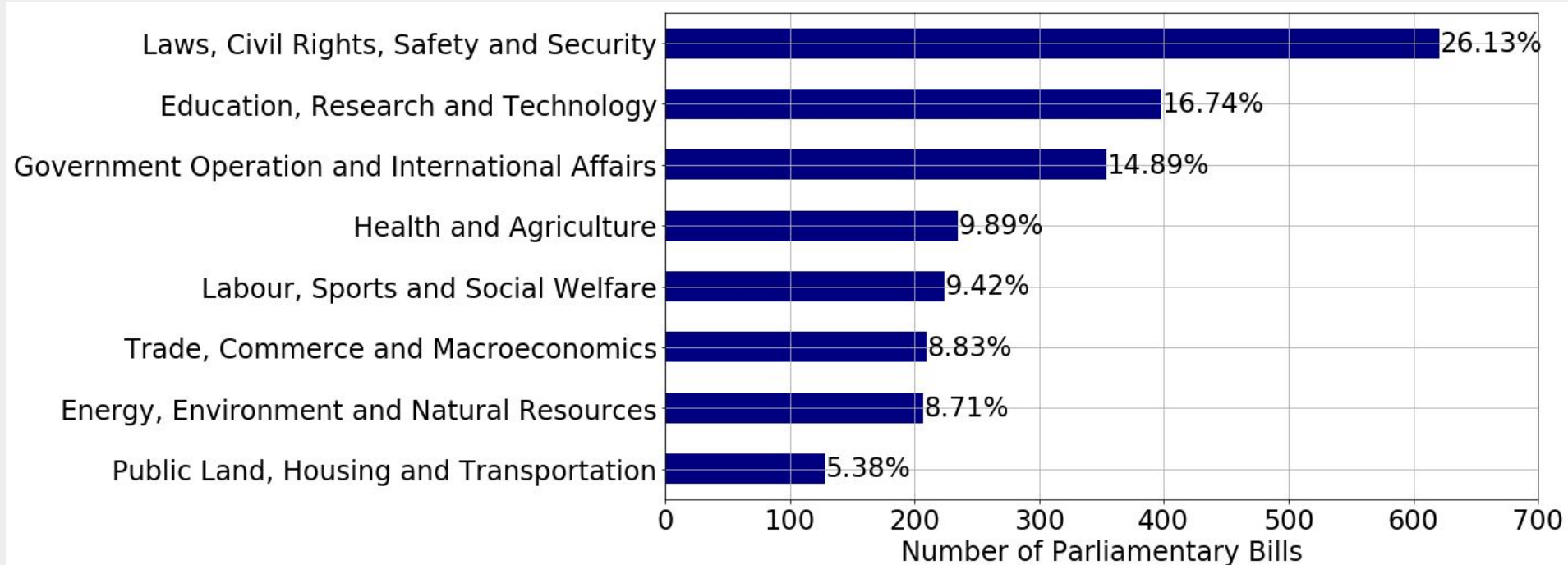
Designated
Hospitals.

Duties and
Obligation of
Medical
officers.

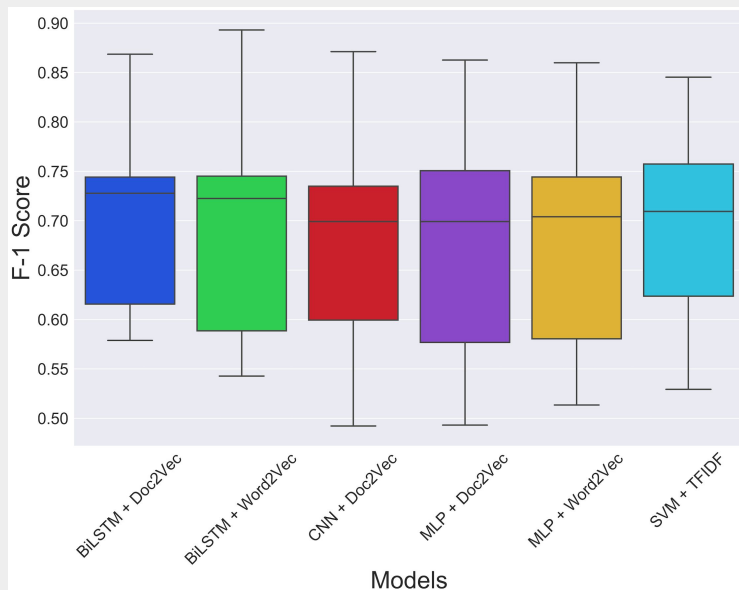
Responsibilities
of the Govern-
ment of the
Federation.

A bill to provide free screening and treatment of cancer and brain tumor

Socio-Economic Driven Labels for Dataset



Results



Models

	Precision	Recall	F1-Score
BiLSTM + Doc2Vec	0.732	0.720	0.718
BiLSTM + Word2Vec	0.718	0.714	0.714
CNN + Doc2Vec	0.711	0.699	0.698
MLP + Doc2Vec	0.710	0.701	0.700
MLP + Word2Vec	0.715	0.701	0.700
SVM + TFIDF	0.722	0.710	0.709

Comparison of BiLSTM + Doc2Vec with other baseline methods

In Proceedings of the NeurIPS 2019 Workshop on Machine Learning for the Developing World (ML4D)

NASS-AI: Towards Digitization of Parliamentary Bills using Document Level Embedding and Bidirectional Long Short-Term Memory

Adewale Akinfaderin*
Data Duality Labs
aaa12g@my.fsu.edu

Olamilekan Wahab*
Independent Researcher
olamy53@gmail.com

Abstract

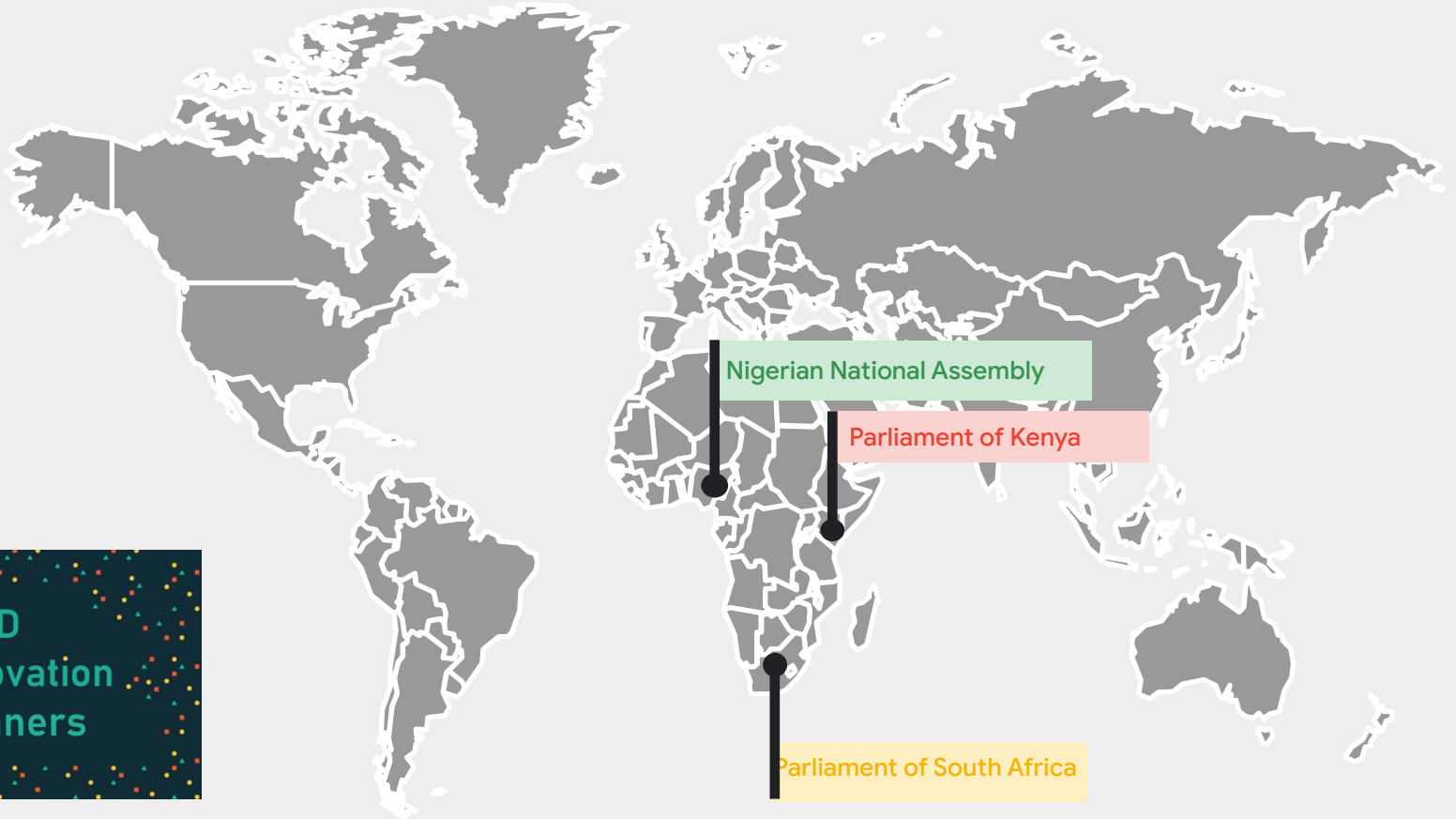
There has been several reports in the Nigerian and International media about the Senators and House of Representative Members of the Nigerian National Assembly (NASS) being the highest paid in the world. Despite this high-level of parliamentary compensation and a lack of oversight, most of the legislative duties like bills introduced and vote proceedings are shrouded in mystery without an open and annotated corpus. In this paper, we present results from ongoing research on the categorization of bills introduced in the Nigerian parliament since the fourth republic (1999 - 2018). For this task, we employed a multi-step approach which involves extracting text from scanned and embedded pdfs with low to medium quality using Optical Character Recognition (OCR) tools and labeling them into eight categories. We investigate the performance of document level embedding for feature representation of the extracted texts before using a Bidirectional Long Short-Term Memory (Bi-LSTM) for our classifier. The performance was further compared with other feature representation and machine learning techniques. We believe that these results are well-positioned to have a substantial impact on the quest to meet the basic open data charter principles.

1 Introduction

Given the challenges and precariousness facing developing and underdeveloped countries, the quality of policy-making and legislation is of enormous importance. This legislation can be used to impact the success of some of the United Nations Sustainable Development Goals (SDGs) like poverty alleviation, good public health system, quality education, economic growth and sustainability. Targets 16.6 and 16.7 from the UN SDGs are to "develop effective, accountable, and transparent institutions at all levels" and to "ensure responsive, inclusive, participatory and representative decision making at all levels" [1]. For countries in Sub-Saharan Africa to meet this target, an open data revolution needs to happen at all levels of government and more importantly, at the parliamentary level.

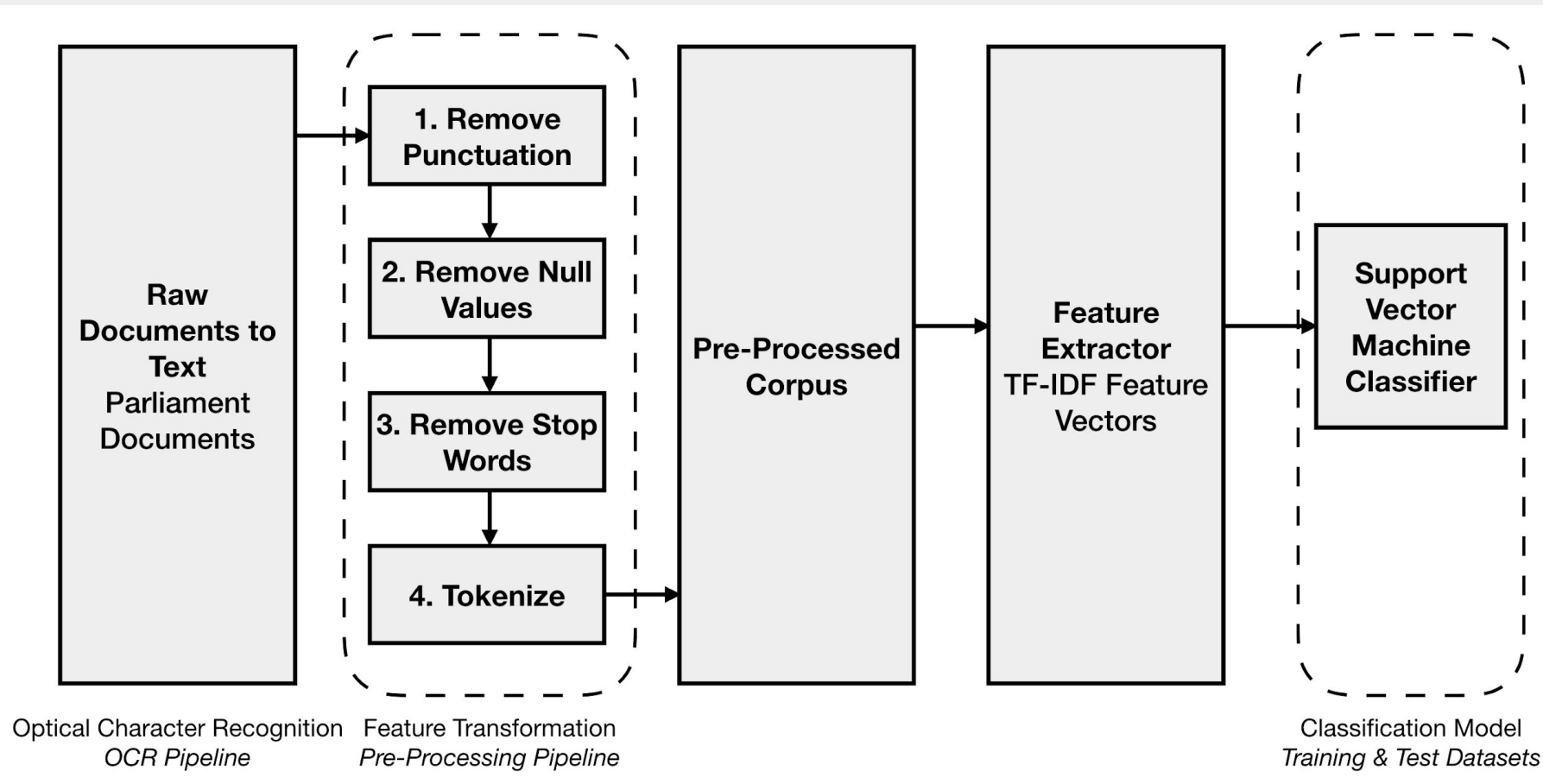
arXiv:1910.04865v1 [cs.CV] 2 Oct 2019

Extend to Other African Countries and Custom OCR



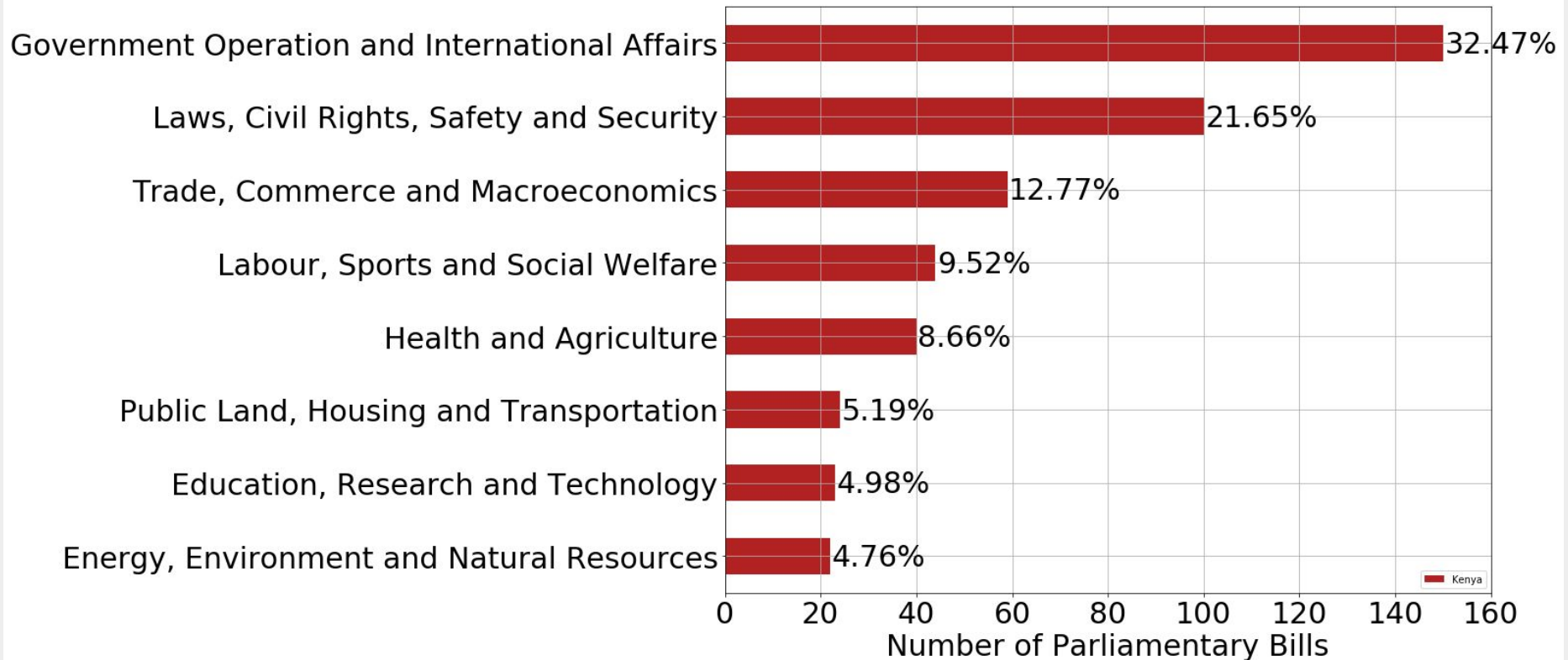
#AI4D
Africa Innovation
2019 Winners

Current Pipeline for Baseline



Kenya Bills

- ~ 460 Bills from 2009 - 2019



Kenya Parliament - Results

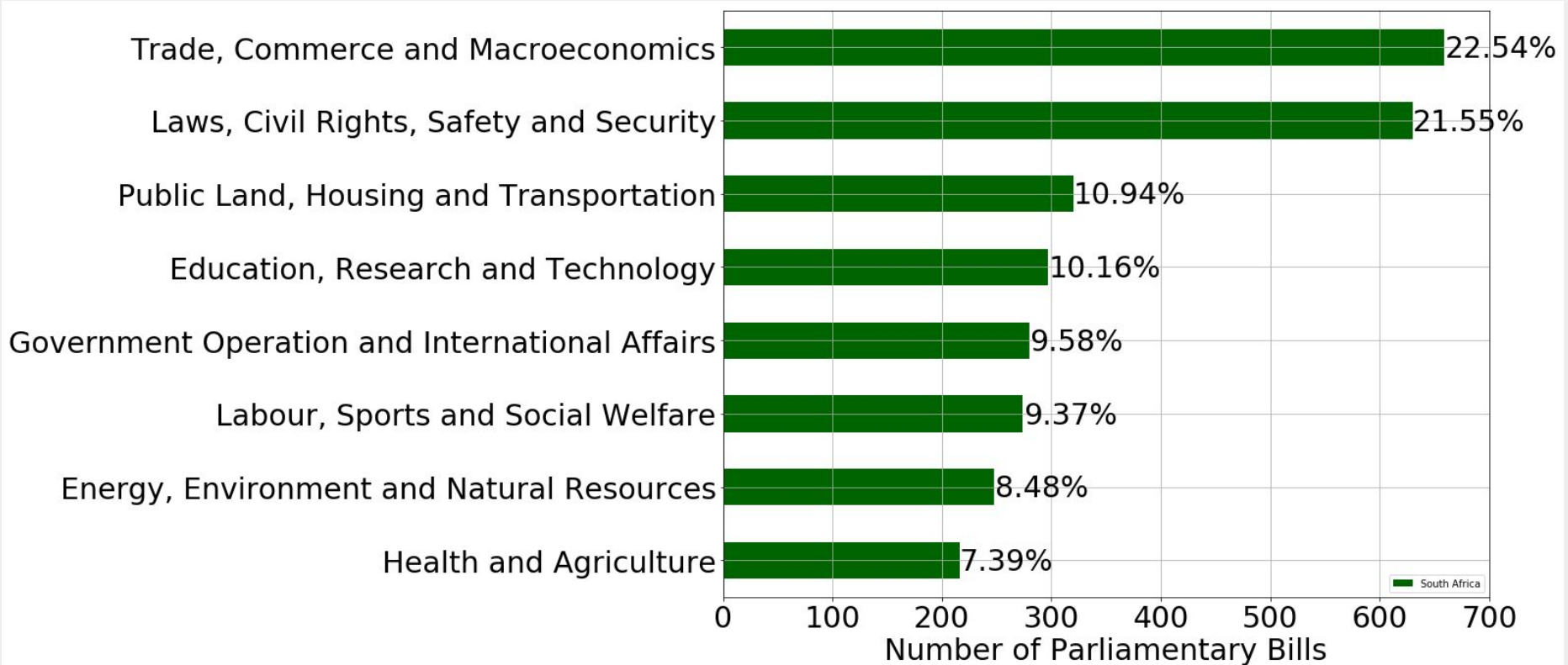
- 70/30 Split with Five Fold CV
- TFIDF + SVM: Average F1 ~ 0.60
- Doc2Vec + SVM: Average F1 ~ **0.87**

F1 score: 0.597702771990628				
	precision	recall	f1-score	support
government operation and international affairs	0.51	0.80	0.62	45
laws, civil rights, safety and security	0.56	0.47	0.51	30
labour, sports and social welfare	0.50	0.31	0.38	13
trade, commerce and macro economics	1.00	0.56	0.71	18
public land, housing and transportation	0.75	0.43	0.55	7
energy, environment and natural resources	0.75	0.43	0.55	7
education, research and technology	1.00	0.71	0.83	7
health and agriculture	0.73	0.73	0.73	11
accuracy			0.60	138
macro avg	0.72	0.55	0.61	138
weighted avg	0.65	0.60	0.60	138

F1 score: 0.8685760354567129				
	precision	recall	f1-score	support
government operation and international affairs	0.81	0.87	0.84	45
laws, civil rights, safety and security	0.84	0.87	0.85	30
labour, sports and social welfare	0.90	0.69	0.78	13
trade, commerce and macro economics	0.89	0.94	0.92	18
public land, housing and transportation	1.00	0.71	0.83	7
energy, environment and natural resources	0.86	0.86	0.86	7
education, research and technology	1.00	1.00	1.00	7
health and agriculture	1.00	1.00	1.00	11
accuracy			0.87	138
macro avg	0.91	0.87	0.89	138
weighted avg	0.87	0.87	0.87	138

South Africa Bills

- ~3000 bills from 1996 - 2019.
- Condensed 98 distinct categories and mapped to our labels (Eight).



South African Parliament - Results

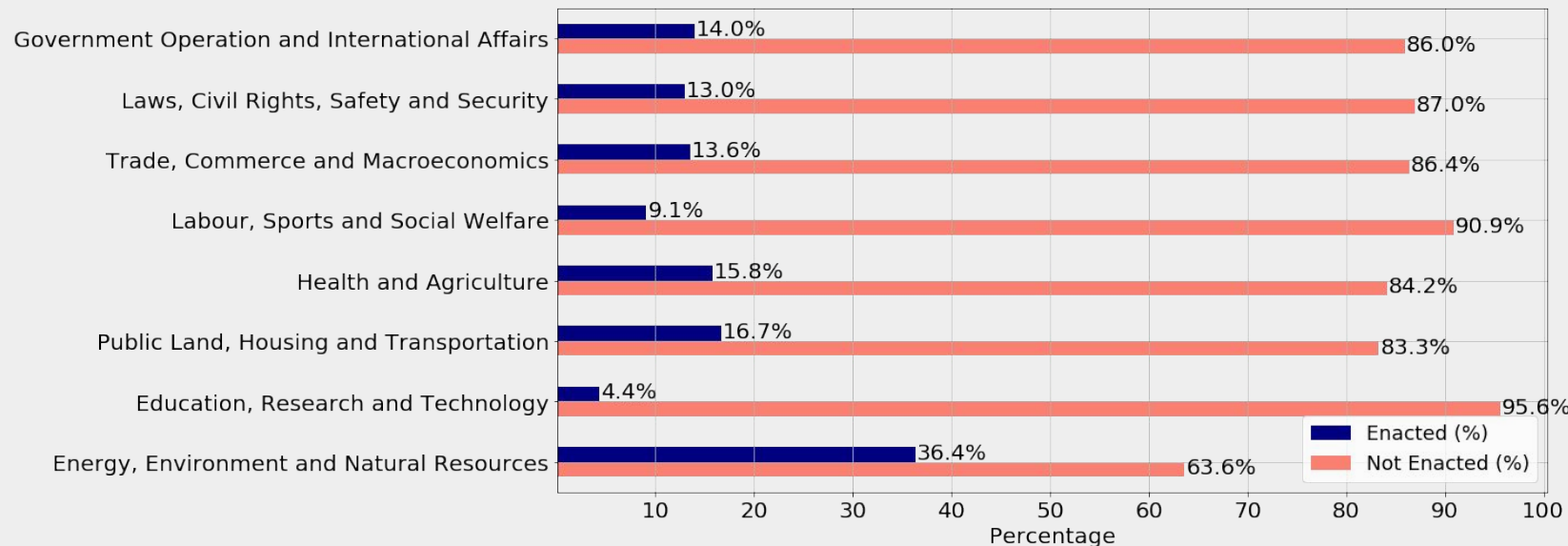
- ~ 3000 Bills from 1996 - 2019
- 70/30 Train-Test split with five fold CV
- Doc2Vec + SVM: Average F1 ~ **0.94**

F1 score: 0.9396036596504206

	precision	recall	f1-score	support
laws, civil rights, safety and security	0.95	0.97	0.96	182
labour, sports and social welfare	0.84	0.94	0.89	86
public land, housing and transportation	0.98	0.95	0.96	95
trade, commerce and macroeconomics	0.92	0.93	0.93	192
education, research and technology	0.95	0.95	0.95	91
energy, environment and natural resources	0.99	0.94	0.96	77
government operation and international affairs	0.94	0.97	0.96	88
health and agriculture	1.00	0.82	0.90	66
accuracy			0.94	877
macro avg	0.95	0.93	0.94	877
weighted avg	0.94	0.94	0.94	877

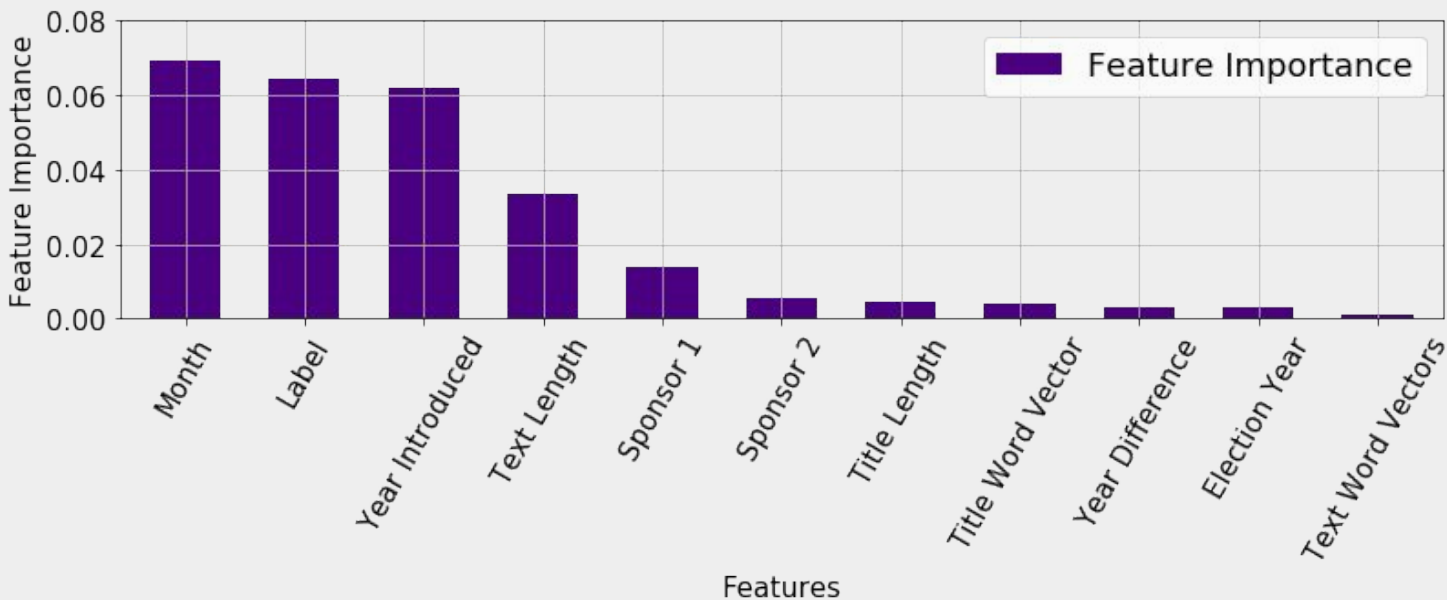
Ongoing: Prediction and Analysis of Law-Making in Kenya*

- Out of 460 bills introduced to the Kenyan parliament from 2009 - 2019, only 65 (14.1%) were enacted.
- Using a combination of handcrafted and text-based features, we developed machine learning algorithm to predict the probability that a bill will become law or not.

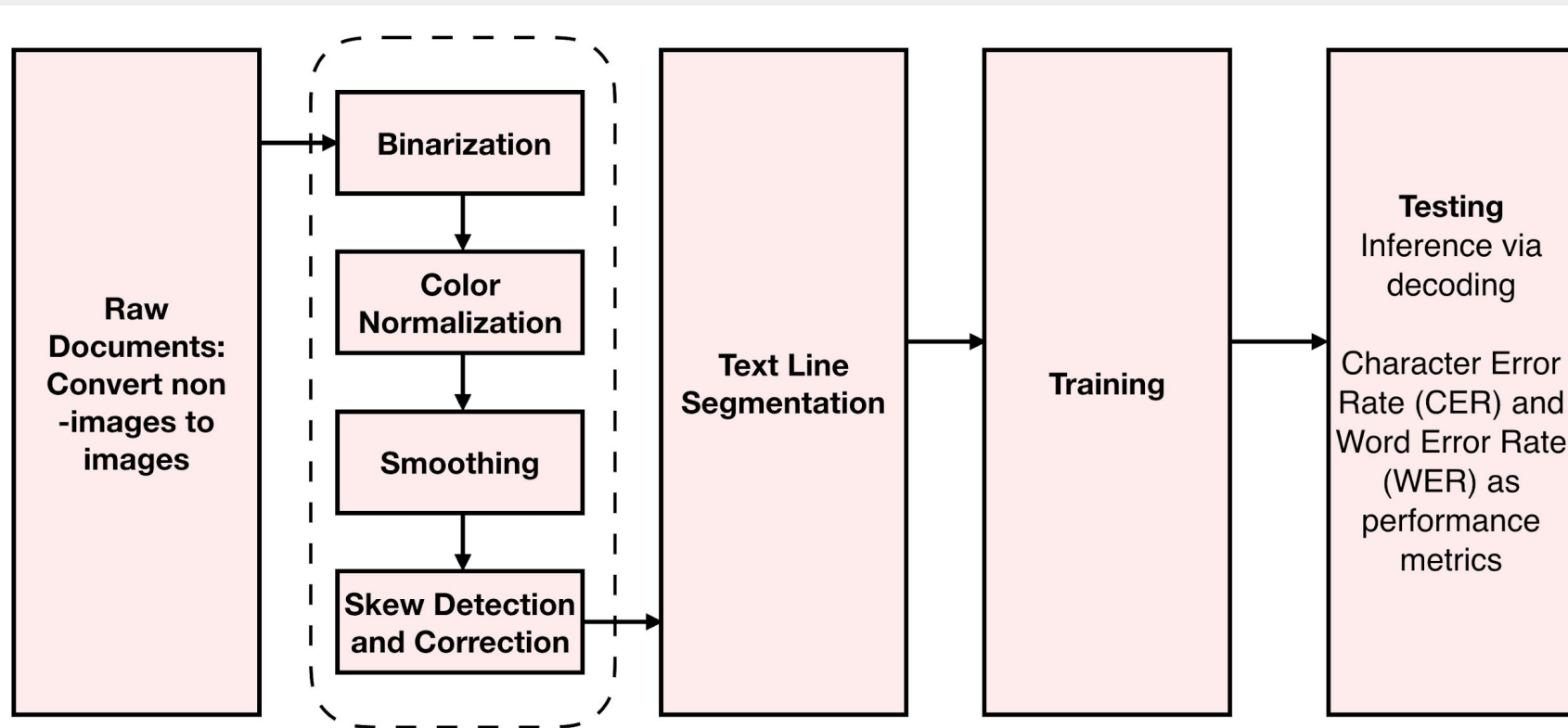


Ongoing: Prediction and Analysis of Law-Making in Kenya*

- Data Imbalance Problem: Synthetic Minority Oversampling Technique (SMOTE)
- Three Models: Logistic Regression, Random Forest and a Stacked Ensemble Model
- Evaluation Metrics: F1 and Brier Score
- Stacked Ensemble Results: Average-F1~0.68, Brier Score~0.39 (Lower is better).



Ongoing: Custom OCR for Low to Medium Quality Bills



Preprocessing with OpenCV
Pre-Processing Pipeline

Future Work: Temporal-based Embedded Topic Modelling for Parliamentary Bills

- The parliamentary bills are collected over a large number of years for most countries
- We plan to focus on analyzing the temporal evolution of topics of the bills and empirically evaluate the changes in the latent patterns of the documents over time.
- To achieve this, we plan to leverage on The Dynamic Embedded Topic Model (D-ETM) recently developed by Dieng et al. *
- D-ETM is a generative model of documents that leverages on word embeddings combined with Dynamic Latent Dirichlet Allocation (D-LDA).

Team



Wale Akinfaderin
waleakinfaderin@gmail.com
Principal Investigator



Oyin Babafemi
babsoyinlola@gmail.com
Machine Learning Engineer

Acknowledgement

- AI4D Innovation Grant
- K4A Foundation
- Data Science Nigeria Team
- Olamilekan Wahab
- Ahmed Baruwa

**I'm hiring an intern for this project. Reach out if interested.*